

**Michael Stephen Saxon**  
[saxon@ucsb.edu](mailto:saxon@ucsb.edu)    <https://saxon.me/>

### Education

**University of California, Santa Barbara**  
*Ph.D.*, Computer Science: **4.0/4.0**  
*Advisors: William Yang Wang, Ph.D.*

Santa Barbara, CA  
9/2020–present

**Arizona State University**  
*MS.*, Computer Engineering: **3.9/4.0**  
*Advisors: Visar Berisha, Ph.D. & Sethuraman Panchanathan, Ph.D.*

Tempe, AZ  
8/2018–5/2020

**Arizona State University**  
*BSE.*, Electrical Engineering; *Minor*, Mathematics: **Magna Cum Laude**

Tempe, AZ  
8/2014–8/2018

### Publications


**Archival** (and publicly available, to-be-archival preprints)

○ Mentor    ^ Representative    ☆ Award

- [c23] **M. Saxon**, A. Holtzman, P. West, WY. Wang, N. Saphra, “Benchmarks as Microscopes: A Call for  
^ Model Metrology,” [saxon.me/doc/metrology.pdf](https://saxon.me/doc/metrology.pdf) COLM 2024
- [ur] **M. Saxon\***, A. Sharma\*, WY. Wang, “Losing Visual Needles in Image Haystacks: Vision Language  
Models are Easily Distracted in Short and Long Contexts,” *Under review*, 2024
- [ur] Q. Wu, H. Zhao, **M. Saxon**, T. Bui, WY. Wang, Y. Zhang, S. Chang, “VSP: Assessing the dual chal-  
lenges of perception and reasoning in spatial planning tasks for VLMs,” *Under review*, 2024
- [ur] E. Tanwar, A. Chatterjee, **M. Saxon**, A. Albalak, WY. Wang, T. Chakraborty, “Do You Know About My  
Nation? An Investigation into Cultural Literacy of Multilingual Large Language Models,” *Under  
review*, 2024
- [ur] W. Feng, J. Li, **M. Saxon**, T. Fu, W. Chen, WY. Wang, “TC-Bench: Benchmarking Temporal Com-  
positionality in Text-to-Video and Image-to-Video Generation,” *Under review*, 2024
- [c22] **M. Saxon\***, Y. Luo\*, S. Levy, C. Baral, Y. Yang, WY. Wang, “Lost in Translation? Translation Errors  
and Challenges for Fair Assessment of Text-to-Image Models on Multilingual Concepts,” **NAACL  
2024**, *Oral (5% of subs.)* [openreview:4uDzjx\\_QX5R0](#) June 2024.
- [p4] **M. Saxon\***, M. Khoshnoodi\*, F. Jahara\*, Y. Lu, A. Sharma, WY. Wang, “Who Evaluates the Eval-  
uations? Assessing the Faithfulness and Consistency of Text-to-Image Evaluation Metrics with  
T2IScoreScore,” [arxiv:2403.11092](#) *Under review*, Feb 2024
- [J21] L. Pan, **M. Saxon**, W. Xu, D. Nathani, X. Wang, WY. Wang, “Automatically Correcting Large Lan-  
guage Models: Surveying the landscape of diverse self-correction strategies,” [arXiv:2308.03188](#),  
**Trans. of the ACL (TACL)** May 2024.
- [c20] V. Himakunthala\*, A. Ouyang\*, D. Rose\*, R. He\*, A. Mei, Y. Lu, C. Sonar, **M. Saxon**, WY. Wang,  
○ “Let’s Think Frame by Frame with VIP: A Video Infilling and Prediction Dataset for Evaluating  
Video Chain-of-Thought,” **EMNLP 2023**, [arXiv:2305.13903](#), Dec 2023
- [c19] X. Wang, W. Zhu, **M. Saxon**, M. Steyvers, WY. Wang, “Large Language Models Are Implicitly Topic  
Models: Explaining and Finding Good Demonstrations for In-Context Learning,” **NeurIPS 2023**,  
[arXiv:2301.11916](#), Dec 2023
- [c18] **M. Saxon**, WY. Wang, “Multilingual Conceptual Coverage in Text-to-Image Models,” **ACL 2023**;  
^ **FAccT 2023 Oral** [arXiv:2306.01735](#), [[oral presentation link](#)] Jul 2023.

- [c17] Y. Tuan, A. Albalak, W. Xu, **M. Saxon**, C. Pryor, L. Getoor, WY. Wang, "CausalDialogue: Modeling Utterance-level Causality in Conversations," **ACL 2023 F** [arXiv:2212.10515](#), Jul 2023.
- [p2] D. Rose\*, V. Himakunthala\*, A. Ouyang\*, R. He\*, A. Mei, Y. Lu, **M. Saxon**, C. Sonar, D. Mirza, WY. Wang, "Visual Chain of Thought: Bridging Logical Gaps with Multimodal Infillings," *preprint*, [arXiv:2305.02317](#), May 2023.
- [p1] **M. Saxon\***, A. Mei\*, S. Chang, ZC. Lipton, WY. Wang, "Users are the North Star for AI Transparency," *preprint*, [arXiv:2303.05500](#), Mar 2023.
- [c16] **M. Saxon**, X. Wang, W. Xu, WY. Wang, "PECO: Examining Single Sentence Label Leakage in Natural Language Inference Datasets," **EACL 2023** [arXiv:2112.09237](#), May 2023.
- [c15] M. Ho\*, A. Sharma\*, J. Chang\*, **M. Saxon**, S. Levy, Y. Lu, WY. Wang, "WikiWhy: Answering and Explaining Cause-and-Effect Questions," **ICLR 2023** *Oral (top 5%.)* [arXiv:2210.12152](#), May 2023.
- [c14] X. Wang, **M. Saxon**, J. Li, H. Zhang, K. Zhang, WY. Wang, "Causal Balancing for Domain Generalization," **ICLR 2023** [arXiv:2206.05263](#), May 2023.
- [c13] W. Xu, Y. Tuan, Y. Lu, **M. Saxon**, L. Li, WY. Wang, "Not All Errors are Equal: Learning Text Generation Metrics using Stratified Error Synthesis," **EMNLP 2022 F** [arXiv:2210.05035](#), Dec 2022.
- [c12] W. Xu, **M. Saxon**, M. Sra, WY. Wang, "Self-Supervised Knowledge Assimilation for Expert-Layman Style Transfer," **AAAI 2022** [arXiv:2110.02950](#), Jan 2022.
- [c11] X. Wang, W. Chen, **M. Saxon**, WY. Wang, "Counterfactual Maximum Likelihood Estimation for Training Deep Networks," **NeurIPS 2021** [arXiv:2106.03831](#), Dec 2021.
- [c10] **M. Saxon**, S. Levy, X. Wang, A. Albalak, WY. Wang, "Modeling Disclosive Transparency in NLP Application Descriptions," **EMNLP 2021** *Oral (8% of subs.)* [arXiv:2101.00433](#), pp. 2023–2037.
- [c9] **M. Saxon**, S. Choudhary, J. McKenna, A. Mouchtaris, "End-to-End Spoken Language Understanding for Generalized Voice Assistants," **Interspeech 2021**, pp. 4738–4742.
- [c8] S. Levy, **M. Saxon**, WY. Wang, "The Truth is Out There: Investigating Conspiracy Theories in Text Generation," [arXiv:2101.00379](#), **Findings of ACL 2021**, pp. 4718–4729.
- [j7] **M. Saxon**, A. Tripathi, Y. Jiao, J. Liss, V. Berisha, "Robust Estimation of Hypernasality in Dysarthria," **IEEE Trans. on Audio, Speech, and Language Processing** 2020, Vol. 28, pp. 2511–2522.
- [c6] **M. Saxon\***, J. McKenna\*, S. Choudhary\*, G. Strimel, A. Mouchtaris, "Semantic Complexity in End-to-End Spoken Language Understanding," **Interspeech 2020**, pp. 4273–4277.
- [c5] M. Moore, P. Papreja, **M. Saxon**, V. Berisha, S. Panchanathan, "UncommonVoice: A Crowdsourced Dataset of Dysphonic Speech," **Interspeech 2020**, pp. 2532–2536.
- [c4] M. Moore, **M. Saxon**, H. Venkateswara, V. Berisha, S. Panchanathan, "Say what? A dataset for exploring the error patterns that two ASR engines make," **Interspeech 2019**, pp. 2528–2532.
- [c3] **M. Saxon**, J. Liss, V. Berisha, "Objective Measures of Plosive Nasalization in Hypernasal Speech," 2019 **IEEE ICASSP 2019**, pp. 6520–6524.
- [w2] **M. Saxon\***, S. Bhandari\*, L. Ruskin, G. Honda, "Word Pair Convolutional Model for Happy Moment Classification," **2<sup>nd</sup> Workshop on Affective Content Analysis**, **AAAI 2019**, pp. 111–119. (*Workshop Oral; CL-Aff Shared task runner up, 2/47*)
- [c1] T. Houghton, **M. Saxon**, Z. Song, H. Nyugen, H. Jiang and H. Yu, "2D Grating Pitch Mapping of a through Silicon Via (TSV) and Solder Ball Interconnect Region Using Laser Diffraction" **IEEE 66th Electronic Components and Technology Conference (ECTC) 2016**, pp. 2222–2227. (*Texas Instruments Best Student Interactive Paper Award*)

## Select Non-archival Presentations

- [n4] **M. Saxon**, WY. Wang, “Disparities in Text-to-Image Model Concept Possession Across Languages,”  **FAccT 2023 Oral** (Non-archival) [OpenReview:5H2m3tCEaQ](#), Jun 2023.
- [n3] **M. Saxon**, X. Wang, W. Xu, WY. Wang, “Automated Cheating Feature Semantic Identification for NLI Datasets,” **SoCal NLP 2022**, Nov 2022.

## Professional Experience

<b>AMD</b> (Open Source Generative AI) <i>Research Intern</i>	Bellevue, WA 6/2024–10/2024
<b>Meta</b> (Facebook Conversational AI) <i>Research Intern</i>	Menlo Park, CA 6/2022–10/2022
<b>Amazon</b> (Alexa Web-based Question Answering) <i>Applied Science Intern</i>	Manhattan Beach, CA 6/2021–9/2021
<b>Amazon</b> (Alexa Edge ML) <i>Applied Science Intern</i> (2x)	Pittsburgh, PA 5/2019–8/2019, 1/2020–8/2020

## Invited Talks

<b>University of Maryland, College Park</b> <i>UMD CLIP Seminar</i>	5/2024
<b>Georgetown University</b> <i>NLP Group Presentation</i>	5/2024
<b>University of Maryland, Baltimore</b> <i>Perception, Prediction, and Reasoning Seminar</i>	4/2024
<b>Arizona State University</b> <i>Active Perception Group Presentation</i>	11/2023
<b>USC Information Sciences Institute</b> <i>Natural Language Processing Seminar</i>	11/2023

## Service

<i>Program Co-Chair</i> , 2022 Southern California NLP Workshop (SoCalNLP)	11/2022
<i>Reviewer</i> , AAAI, EMNLP, ACL, EACL, NeurIPS, ICLR, ACM FAccT, ICASSP, Interspeech	2020–present
<i>Mentor</i> , FIRST Robotics Team 2478 (Westwood Robotics), Mesa, AZ	2014–2016

## Mentoring

Mahsa Khoshnoodi	2023–2024	Fatima Fellowship Mentee	⇒	PhD. admit, <b>Stony Brook</b>
Fatima Jahara	2023–2024	Fatima Fellowship Mentee	⇒	PhD. admit, <b>Rutgers</b>
Matthew Ho	2021–2024	UCSB Undergrad	⇒	Ph.D. admit, <b>UC San Diego</b>
Aditya Sharma	2021–2024	UCSB Undergrad, MS	⇒	Incoming SWE, <b>Google</b>
Avani Tanna	2022–2023	UCSB MS student	⇒	GenAI Engineer, <b>Domino.AI</b>
Justin Chang	2021–2023	UCSB Undergrad	⇒	SWE, <b>Aptos</b>

## Honors, Fellowships, Scholarships

<b>University of California, Santa Barbara</b> <i>Neal Fenzi—Resonant Founder Fellowship</i>	2024
<b>National Science Foundation</b> <i>Graduate Research Fellowship</i> (NSF GRFP)	2020
<b>University of California, Santa Barbara</b> <i>Center for Responsible Machine Learning Fellowship</i>	2020
<b>University of California, Santa Barbara</b> <i>Graduate Division Central Fellowship</i>	2020
<b>Arizona State University</b> <i>Presidential Scholarship</i> (Full Tuition)	2014